NEW YORK UNIVERSITY
Institute of Mathematical Sciences
Division of Electromagnetic Research

RESEARCH REPORT No. BR-8

A Fourier Theorem for Matrices

WILHELM MAGNUS

CONTRACT No. AF-18(600)-367

OCTOBER 1954

NEW YORK UNIVERSITY

Institute of Mathematical Sciences

Division of Electromagnetic Research

Research Report No. BR-8

A FOURIER THEOREM FOR MATRICES

by

Wilhelm Magnus

Wilhelm Magnus
Wilhelm Magnus

Morris Kline
Morris Kline
Project Director

The research reported in this document has been made possible through support and sponsorship extended by the Office of Scientific Research, H.Q., Air Research and Development Command, U.S.A.F., Baltimore, Maryland, under Contract No. AF-18(600)-367. It is published for technical information only, and does not necessarily represent recommendations or conclusions of the sponsoring agency.

New York, 1954

Abstract

A Fourier theorem is proved which establishes a one-to-one correspondence between $n \times n$ matrices depending on a real parameter and $n \times n$ matrices depending on the elements of a variable hermitian matrix and satisfying certain differential equations. The analogue of the Plancherel formula is shown to be true for this Fourier theorem.

Table of Contents

1. Introduction and Summary

In a previous report[1] a study of the Baker-Hausdorff formula and its application to systems of linear differential equations was made. The results given in the present paper are based on those of reference [1].

We start by considering the nature of the mapping

$$(1.1) \qquad U = \exp(i H),$$

where H is a hermitian matrix of n rows and columns, and where consequently U is a unitary matrix. We may consider H and U as points in a space of n^2 real dimensions, and then (1.1) defines a mapping of one of these spaces upon the other. As coordinates in the space S_H of matrices H we may choose the real and imaginary parts of the elements of H. If

$$(1.2) \qquad H = (k_{\nu,\mu}) = (s_{\nu,\mu} + i a_{\nu,\mu}), \quad \nu,\mu = 1, \ldots, n,$$

$$s_{\nu,\mu} = s_{\mu,\nu}, \quad a_{\nu,\mu} = -a_{\mu,\nu},$$

then the values of the variables $s_{\nu,\mu}$ and $a_{\nu,\mu}$ determine H uniquely, and vice versa. S_H can also be considered as a vector space, since any linear combination of hermitian matrices with constant real coefficients is a hermitian matrix again. The neighborhood of a point in S_H can be defined in a natural way by introducing the Euclidean distance between two points whose Cartesian coordinates are the $a_{\nu,\mu}$, $s_{\nu,\mu}$.

Since the unitary matrices U form a multiplicative group, the natural definition of a neighborhood of a point in the space S_U of matrices U must be derived from a definition of a neighborhood of the identity I. We shall say that U is in a neighborhood of the matrix U_o if $U U_o^{-1}$ is in the neighborhood of I. A neighborhood of I is defined by all those unitary matrices V for which

$$(1.3) \qquad \sum_{\nu,\mu=1}^{n} |u_{\nu,\mu} - \delta_{\nu,\mu}|^2 \leqq \varepsilon^2 , \quad V = (u_{\nu,\mu}), \ I = (\delta_{\nu,\mu}) .$$

The manifold S_U of matrices U is a part of the linear space of all matrices M; this space has $2n^2$ real dimensions. Since every U can be expressed by (1) in terms of a matrix H, we may introduce the $s_{\nu,\mu}$, $a_{\nu,\mu}$ as coordinates in S_U. In terms of these coordinates we shall define in S_U the volume element $d\tau$, which has the property of being invariant under the multiplicative group. This means that the volume of a small region R in the neighborhood of a point U_o on S_U will be measured in terms of the volume of the region $R\,U_o^{-1}$ in the neighborhood of I. Here $R\,U_o^{-1}$ is defined as the set of points or unitary matrices on S_U obtained from the set of matrices belonging to R by right multiplication by U_o^{-1}. The result is

$$(1.4) \qquad d\tau = \prod_{\nu<\mu} \left\{ \frac{2\sin\frac{1}{2}(\lambda_\nu - \lambda_\mu)}{\lambda_\nu - \lambda_\mu} \right\}^2 \prod_{\nu \leqq \mu} ds_{\nu,\mu} \prod_{\nu<\mu} da_{\nu,\mu},$$

where the $\lambda_\nu (\nu = 1,\dots, n)$ are the eigenvalues of H and where all products are to be taken over $\nu,\mu = 1,\dots, n$ with the restrictions indicated below the $\prod$- symbols.

It should be observed that the first product on the right-hand side of (1.4) is an entire symmetric function of the λ_ν and therefore can be expressed as an entire function of the coefficients of the characteristic equation of H. It becomes unity if H is the null matrix.

Any other invariant volume element can differ from $d\tau$ only by a factor which is independent of H.

Now we have the

Lemma 1:

Let S denote the region in S_H which is defined by

$$(1.5) \quad |\lambda_\nu - \lambda_\mu| \leqq 2\pi \qquad (\nu,\mu = 1,\dots, n).$$

Then S is the largest connected region of S_H which contains the null matrix $H = 0$ and which is such that a full neighborhood of any interior point H_o of S is mapped upon a full neighborhood of $U_o = \exp(i H_o)$ in S_U. This region S is needed for the following

Fourier Theorem:

Let

$$F(t) = \left(f_{\nu,\mu}(t)\right) , \qquad (\nu,\mu = 1,\dots, n)$$

be a matrix whose elements $f_{\nu,\mu}$ depend on a parameter t; suppose also that F is defined for $-\infty < t < \infty$ and that

$$(1.7) \quad \int_{-\infty}^{+\infty} |f_{\nu,\mu}(t)| \, dt < \infty \qquad (\nu,\mu = 1,\dots, n).$$

Let H be any hermitian matrix represented by a point in S. Then

$$(1.8) \quad \int_{-\infty}^{\infty} F(t) \exp(it H) \, dt = G(H)$$

exists, and

$$(1.9) \quad \iint_S G(H) \exp(-itH) \, d\tau = L_n F(t).$$

where the scalar L_n depends on n but not on H or F.

If the integral in (1.9) does not converge absolutely, it may be necessary to prescribe an appropriate method of evaluation. As a supplement to the Fourier theorem we have the

Plancherel theorem:

If the elements $f_{\nu,\mu}$ are L^2, then

$$(1.10) \quad \text{trace} \int_{-\infty}^{\infty} F^*(t) F(t) \, dt = L_n^{-1} \, \text{trace} \iint_S G^*(H) G(H) \, d\tau.$$

where the asterisk denotes the complex conjugate of the transpose of a matrix.

Some of the properties of the matrices G(H) which arise from relation (1.8) will be given in Section 5. We shall show there that the elements of G are linear combinations of partial derivatives of unitary invariants of H (the partial derivatives are taken with respect to the elements of H). In Section 5 we also define these unitary invariants and give the partial differential equations by which these unitary invariants can be determined.

2. Computation of the Volume Element

Consider a matrix H + dH, where

$$(2.1)\qquad dH = (ds_{\nu,\mu}) + i(da_{\nu,\mu}), \quad (\nu,\mu = 1,\dots, n),$$

$$ds_{\nu,\mu} = ds_{\mu,\nu}, \quad da_{\nu,\mu} = - da_{\mu,\nu}$$

Now we proceed to compute the quantity

$$(2.2)\qquad \exp(iH + idH) \quad \exp(-iH).$$

The terms in (2.2) that are linear in dH are given by

$$(2.3)\qquad idV = idH - \frac{1}{2!}\,[idH, iH] + \frac{1}{3!}\Big[\,[idH, iH]\; iH\Big] \mp \dots$$

(see [1]), where, for any matrices A, B,

$$(2.4)\qquad [A,B] = AB - BA, \qquad \Big[\,[A,B]B\Big] = AB^2 - 2BAB + B^2A\;, \;\dots$$

The matrix dV is hermitian; if we write

$$(2.5)\qquad dV = (d\,\sigma_{\nu,\mu}) + i(d\alpha_{\nu,\mu}) \qquad (\nu,\mu = 1,\dots,n)$$

then the n^2 variables $d\sigma_{\nu,\mu}(\nu \leq \mu)$ and $d\alpha_{\nu,\mu}(\nu < \mu)$ become linear functions of the n^2 variables $ds_{\nu,\mu}(\nu \leq \mu)$ and $da_{\nu,\mu}(\nu < \mu)$.

The determinant D of these n^2 linear functions is the factor of $\prod ds_{\nu,\mu} \prod da_{\nu,\mu}$ in equation (1.4) times another factor which is a power of i.

We shall determine D by diagonalizing H. For this purpose we shall first apply the following preliminary consideration. Let w be any unitary matrix of n rows and columns. Let S_H be the space of n^2 real dimensions whose points correspond to the hermitian matrices. We shall represent the points in this space by the variables

$$(2.6) \qquad \begin{aligned} x_{\mu,\nu} &= x_{\nu,\mu} \qquad (\nu \leqq \mu), \\ y_{\nu,\mu} &= -y_{\mu,\nu} \qquad (\nu < \mu), \end{aligned} \qquad (\nu,\mu = 1,\ldots,n)$$

where $x_{\nu,\mu}$ may stand for either $s_{\nu,\mu}$ or $d\sigma_{\nu,\mu}$, and $y_{\nu,\mu}$ may stand for either $da_{\nu,\mu}$ or $d\alpha_{\nu,\mu}$. Putting

$$(2.7) \qquad Z = (x_{\nu,\mu} + i\, y_{\nu,\mu})$$

we obtain a linear mapping of S_H upon itself by choosing an arbitrary unitary matrix W and putting

$$(2.8) \qquad W^{-1} Z W = Z' = (x'_{\nu,\mu} + i\, y'_{\nu,\mu}),$$

where the $x'_{\nu,\mu}$, $y'_{\nu,\mu}$ are homogeneous linear functions of the $x_{\nu,\mu}$, $y_{\nu,\mu}$. Now, we claim that the determinant of the coefficients of these n^2 functions of n^2 variables equals unity. Since the coefficients of the linear functions are real, the determinant must be real. Now, from $W^{-1} = W^*$ we find

$$(2.9) \qquad Z' Z'^* = W^* Z W W^* Z^* W = W^{-1} Z Z^* W.$$

Therefore,

$$(2.10) \qquad \text{trace } Z' Z'^* = \text{trace } Z Z^*,$$

which shows that the transformation of S_H which maps Z upon Z' is an orthogonal one. Its determinant is therefore ± 1. It is certainly + 1 if W is the identity. In fact, it is always unity, since the determinant is a continuous function of the elements of W, and since the space of unitary matrices is connected.

Now we choose a unitary matrix W such that

$$(2.11) \qquad W^{-1} H W = \left\{ \lambda_1, \ldots\ldots, \lambda_n \right\} = \Lambda$$

is a diagonal matrix with the eigenvalues $\lambda_1, \ldots, \lambda_n$ of H in the main diagonal. Of course, the elements of W depend on the elements of H, but it is known that a unitary matrix W satisfying (2.11) always exists.

Introducing

$$(2.12) \qquad d\hat{V} = W^{-1}(dV)W, \quad d\hat{H} = W^{-1}(dH)W .$$

we obtain from (3)

$$(2.13) \qquad id\hat{V} = id\hat{H} - \frac{1}{2!}\left[id\hat{H}, i\Lambda\right] + \frac{1}{3!}\left[\left[id\hat{H}, i\Lambda\right] i\Lambda\right] \mp \ldots \quad .$$

According to (2.12), the coordinates or variables defining $d\hat{V}$, $d\hat{H}$ are obtained by one and the same orthogonal transformation from those defining $d\hat{H}$. Consequently, the determinant of the linear substitution (2.13) connecting $d\hat{V}$ and $d\hat{H}$ has the value D, which is the value of the determinant of the linear substitution connecting dV and dH. Putting

$$(2.14) \qquad \begin{aligned} d\hat{V} &= (d\hat{\sigma}_{\nu,\mu}) + i(d\hat{\alpha}_{\nu,\mu}) \\ d\hat{H} &= (d\hat{s}_{\nu,\mu}) + i(d\hat{a}_{\nu,\mu}) , \end{aligned}$$

where $d\hat{\sigma}_{\nu,\mu} = d\hat{\sigma}_{\mu,\nu}$, $d\hat{\alpha}_{\nu,\mu} = - d\hat{\alpha}_{\mu,\nu}$ etc., we find immediately from (2.11) and (2.13) that

$$(2.15) \qquad d\hat{\sigma}_{\nu,\nu} = d\hat{s}_{\nu,\nu},$$

$$(2.16) \qquad d\hat{\sigma}_{\nu,\mu} + i d\hat{\alpha}_{\nu,\mu} = (d\hat{s}_{\nu,\mu} + i d\hat{a}_{\nu,\mu}) \frac{\exp\left[i(\lambda_\nu - \lambda_\mu)\right] - 1}{i(\lambda_\nu - \lambda_\mu)}, \qquad \nu < \mu,$$

or

$$(2.17) \qquad \begin{aligned} d\hat{\sigma}_{\nu,\mu} &= d\hat{s}_{\nu,\mu} \frac{\sin(\lambda_\nu - \lambda_\mu)}{\lambda_\nu - \lambda_\mu} + d\hat{a}_{\nu,\mu} \frac{\cos(\lambda_\nu - \lambda_\mu) - 1}{\lambda_\nu - \lambda_\mu} \\ d\hat{\alpha}_{\nu,\mu} &= -d\hat{s}_{\nu,\mu} \frac{\cos(\lambda_\nu - \lambda_\mu) - 1}{\lambda_\nu - \lambda_\mu} + d\hat{a}_{\nu,\mu} \frac{\sin(\lambda_\nu - \lambda_\mu)}{\lambda_\nu - \lambda_\mu} . \end{aligned}$$

The determinant of the linear substitution (2.17) connecting the elements of $d\hat{V}$ and $d\hat{H}$ is clearly

$$(2.18) \qquad D = \prod_{\nu < \mu} \left\{ \frac{2 \sin \frac{1}{2} (\lambda_\nu - \lambda_\mu)}{\lambda_\nu - \lambda_\mu} \right\}^2 .$$

This completes the proof of formula (1.4) for the volume element $d\tau$.

3. The Fourier Theorem

We proceed now with the proof of the Fourier theorem stated in (1.6)-(1.9). We need the following

Lemma 2:†

Let $H = (s_{\nu,\mu} + i\, a_{\nu,\mu})$ be a hermitian matrix. Let

$$(3.1) \qquad \varepsilon_{\nu,\mu} = \frac{1}{2}, \qquad (\nu \neq \mu), \qquad \varepsilon_{\nu,\nu} = 1,$$

and let ∇_H be the matrix differential operator

$$(3.2) \qquad \nabla_H = \left(\varepsilon_{\nu,\mu} \frac{\partial}{\partial s_{\nu,\mu}} + i\, \varepsilon_{\nu,\mu} \frac{\partial}{\partial a_{\nu,\mu}} \right) .$$

---

† I am indebted to Prof. B. Friedman for his simple derivation of (3.3) from (3.7), which is used in the proof of this Lemma.

∇_H is obtained from the matrix H by substituting for each of the elements of H either the operator of partial differentiation with respect to this element, or one-half of this operator, depending on whether the element contains $\varepsilon_{\nu,\mu}$ or $\varepsilon_{\nu,\nu}$. Then we have

$$(3.3) \qquad e^{itH} = \sum_{\nu=1}^{n} e^{i\lambda_\nu t} \nabla_H \lambda_\nu ,$$

where the λ_ν are the eigenvalues of H. Explicitly the matrix $\nabla_H\lambda_\nu$ is defined by

$$\nabla_H \lambda_\nu = \left(\varepsilon_{\nu,\mu} \frac{\partial\lambda}{\partial s_{\nu,\mu}} + i\, \varepsilon_{\nu,\mu} \frac{\partial\lambda}{\partial a_{\nu,\mu}}\right),$$

where we have used the natural definition

$$\frac{\partial\lambda}{\partial s_{\nu,\mu}} = \frac{\partial\lambda}{\partial s_{\mu,\nu}}, \quad \frac{\partial\lambda}{\partial a_{\nu,\mu}} = -\frac{\partial\lambda}{\partial a_{\mu,\nu}} .$$

We shall prove Lemma 2 by using a formula due to Sylvester. Let

$$(3.4) \qquad P(\lambda) = |\lambda I - H|$$

be the characteristic polynomial of H. Its roots are the eigenvalues λ_ν of H, and we put

$$(3.5) \qquad \frac{dP(\lambda)}{d\lambda} = P'(\lambda), \qquad p_\nu = P'(\lambda_\nu) \qquad (\nu = 1,\ldots, n),$$

$$(3.6) \qquad P_\nu(\lambda) = \frac{P(\lambda)}{\lambda-\lambda_\nu} .$$

Then we have

$$(3.7) \qquad e^{itH} = \sum_{\nu=1}^{n} e^{it\lambda_\nu} \frac{P_\nu(H)}{p_\nu},$$

provided that the λ_ν are different from each other; in this case, (3.7) can be proved easily by transforming H into a diagonal matrix. But even if we pass to the

limit and several of the λ_ν become equal, (3.7) makes sense, as we shall prove later from Eq. (3.12).

We shall first prove (3.3) for the case where all the λ_ν are different from each other. We proceed as follows:

Let $\nu = 1,\ldots, n$, and let

(3.8) $$x^{(\nu)} = (x_1^{(\nu)}, \ldots , x_n^{(\nu)})$$

be the set of orthonormal eigenvectors of H such that $x^{(\nu)}$ belongs to λ_ν. We consider $x^{(\nu)}$ as a matrix of one column. The transpose and complex conjugate vector of $x^{(\nu)}$ will be denoted by $x^{(\nu)*}$; it is a matrix of one row. The inner product $(x^{(\nu)*}, x^{(\mu)})$ equals $\delta_{\nu,\mu}$, where $\delta_{\nu,\mu}$ is the Kronecker symbol. If we put:

(3.9) $$P_\nu(H)/p_\nu = H_\nu,$$

then clearly

(3.10) $$H_\nu\, x^{(\mu)} = \delta_{\nu,\mu}\, x^{(\mu)} \quad .$$

The matrix H_ν is uniquely defined by (3.10), since if there were two matrices H_ν and H_ν satisfying (3.10), then their difference G_ν would satisfy

(3.11) $$G_\nu\, x^{(\mu)} = 0$$

for $\mu = 1, 2,\ldots, n$, and this is impossible if $G \neq 0$ because the $x^{(\mu)}$ span the n-dimensional space. Then from the definition (3.10) we know that the element in the j-th row and in the k-th column of H_ν must be

(3.12) $$x_j^{(\nu)}\ \bar{x}_k^{(\nu)} \quad ,$$

where the bar denotes the complex conjugate quantity. (Any matrix having these elements (3.12) satisfies equation (3.10) and hence must be identical with H_ν).

Now we are prepared to prove (3.3). We have

$$(3.13) \qquad (H - \lambda_\nu I)\, x^{(\nu)} = 0.$$

If we differentiate with respect to y, where y stands for one of the variables $s_{\nu,\mu}$, $a_{\nu,\mu}$, we find

$$(3.14) \qquad 0 = \frac{\partial}{\partial y}(H - \lambda_\nu I)\, x^{(\nu)} = (H - \lambda_\nu I)\frac{\partial x^{(\nu)}}{\partial y} + \left(\frac{\partial H}{\partial y} - \frac{\partial \lambda_\nu}{\partial y} I\right) x^{(\nu)}.$$

Multiplying the left side of (3.14) by $x^{(\nu)*}$, we obtain

$$(3.15) \qquad x^{(\nu)*}\,\frac{\partial H}{\partial y}\,x^{(\nu)} = x^{(\nu)*}\,\frac{\partial \lambda_\nu}{\partial y}\,x^{(\nu)} = \frac{\partial \lambda_\nu}{\partial y}\,.$$

Now $\partial H/\partial y$ is a matrix with one or two elements different from zero. If $y = s_{\nu,\nu}$, only one element of $\partial H/\partial y$ equals unity and all the other vanish. If $y = s_{\nu,\mu}$, $\nu \neq \mu$, two of the elements of $\partial H/\partial y$ are unity, and if $y = a_{\nu,\mu}$, one element is $+ i$ and one is $-i$. Computing the left-hand side of (3.15) for each of these cases and using (3.12) as an expression for the elements of H_ν in (3.9) we arrive at (3.3).

From the form (3.12) of the elements of the matrix (3.9) we can derive the following:

<u>Lemma 3</u>:

Let the elements of H depend linearly on a parameter ρ in such a way that the eigenvalues of H are different from each other if ρ is sufficiently small but not equal to 0. Then $\lim_{\rho \to 0} H_\nu$ exists and the moduli of its elements are not greater than unity.

<u>Proof</u>: The elements of the eigenvectors of H are of the form

$$(3.16) \qquad D_k \left\{ \sum_{k=1}^{n} \overline{D}_k\, D_k \right\}^{-1/2},$$

where the D_k are determinants involving the elements of H and its eigenvalues.

All of these are single-valued analytic functions of a fractional power $\rho^{1/\ell}$ (ℓ integral) in the neighborhood of $\rho = 0$. Not all the D_k vanish as $\rho \to 0$, except at $\rho = 0$. Therefore, the limit of the expression (3.16) exists for $\rho \to 0$.

It can be shown that every point in the space S_H of matrices H can be reached by a "straight line" of the type described in Lemma 2. Since the points in S_H on which not all the λ_ν are different from each other form an algebraic manifold, it follows that the elements of the matrix H_ν are integrable bounded functions in S_H.

Now we need a decomposition of S_H into a one-parameter set of manifolds $S(\sigma)$. We proceed as follows.

<u>Definition:</u> Let $S(\sigma)$ be the set of all points in S_H for which

(3.17) $$\text{trace } H = s_{11} + s_{22} + \dots + s_{nn} = n\sigma$$

is a fixed multiple of σ. Then we have:

<u>Lemma 4:</u>

$S(0)$ is a linear subspace of S_H. The transformation

(3.18) $$\hat{H} = H + \sigma I,$$

which maps S_H onto itself by mapping H upon $\hat{H}$, also maps $S(0)$ onto $S(\sigma)$. We can replace the coordinates $s_{11}, \dots, s_{nn}$ in S_H by n linear homogeneous functions $\sigma, \rho_1, \dots, \rho_{n-1}$ of these coordinates; these functions are chosen such that the volume element $d\tau$ in S_H can be written as

(3.19) $$d\tau = \sqrt{n}\, d\tau_o\, d\sigma$$

where

(3.20) $$d\tau_o = D \left\{ \prod_{\nu < \mu} (ds_{\nu,\mu}\, da_{\nu,\mu}) \right\} d\rho_1\, d\rho_2 \cdots d\rho_{n-1} .$$

(Here we used the expression (2.18) for D). The value of D is the same in all points of H which can be mapped upon each other by the transformation (3.18); that is, D does not depend on σ. Then the matrix of the substitution connecting the $s_{11}, \ldots, s_{nn}$ and the variables $\sqrt{n}\,\sigma, \rho_1, \ldots, \rho_{n-1}$ can be chosen to be a real orthogonal matrix.

The proof of Lemma 4 is almost obvious. We choose a vector $\vec{v}_o = \frac{1}{\sqrt{n}}(1, 1, \ldots, 1)$ and n-1 vectors $\vec{v}_1, \ldots, \vec{v}_{n-1}$ which, together with $\vec{v}_o$, form the rows of an orthogonal matrix. Putting $\vec{s} = (s_{11}, \ldots, s_{nn})$ and

$$(3.21) \qquad \sqrt{n}\,\sigma = (\vec{v}_o, \vec{s}), \quad \rho_1 = (\vec{v}_1, \vec{s}), \ldots, \rho_{n-1} = (\vec{v}_{n-1}, \vec{s})$$

we obtain the required transformation of coordinates in S_H and the expression for $d\tau_o$ in (3.19). Now we have merely to show that $\lambda_\nu - \lambda_\mu$ is independent of σ; then the same is true for D. But

$$(3.22) \qquad \lambda_\nu = \lambda_{\nu,o} + \sigma,$$

where $\lambda_{\nu,o}$ is the value derived from λ_ν by keeping fixed all coordinates $s_{\nu,\mu}, a_{\nu,\mu} (\nu < \mu)$ and $\rho_1, \ldots, \rho_{n-1}$ in S_H and replacing σ by zero. This completes the proof of Lemma 4.

<u>Lemma 5:</u>

The set $\bar{S}(\sigma)$ of points in $S(\sigma)$ at which not all of the λ_ν are different from each other is given by σ = constant and an algebraic relation between the coordinates $s_{\nu,\mu}, a_{\nu,\mu} (\nu < \mu)$ and $\rho_1, \ldots, \rho_{n-1}$.

<u>Proof:</u> The discriminant of the algebraic equation for the λ_ν is a polynomial in the coordinates of H_S. It does not vanish identically for any given σ as a function of the $s_{\nu,\mu}, a_{\nu,\mu} (\nu < \mu)$ and $\rho_1, \ldots, \rho_{n-1}$, since hermitian matrices with eigenvalues different from each other can be constructed for any preassigned value of $\sigma = (\lambda_1 + \ldots + \lambda_n)/n$.

Lemma 6:

The elements of the matrix H_ν in (3.9) do not depend on σ if $s_{11}, \ldots, s_{nn}$ are replaced by $\sqrt{n}\ \sigma$ and $\rho_1, \ldots \rho_n$. They are bounded and integrable functions in every finite part of S_H.

Proof: The independence of the elements of H_ν from σ follows from the independence of $\lambda_\nu - \lambda_\mu$ from σ and from (3.9). For all points of H for which the λ_ν are different from each other, (3.9) also guarantees the continuity of the elements of H_ν. The rest follows from Lemma 5 (whose analogue for S_H is also true), and from Lemma 3.

Now we are ready to prove the Fourier theorem. Let

(3.23) $$\int_{-\infty}^{+\infty} F(t)\, e^{i\lambda t}\, dt = B(\lambda).$$

Then we have from (3.3) and (3.7):

(3.24) $$\int_{-\infty}^{+\infty} F(t)\, e^{itH}\, dt = \sum_{\nu=1}^{n} B(\lambda_\nu)\, H_\nu = G(H)\ ,$$

where we used the notation of Section 1. Multiplying the right-hand side of (3.24) by

(3.25) $$e^{-itH} = \sum_{\nu=1}^{n} e^{-i\lambda_\nu t}\, H_\nu$$

and observing that according to (3.12)

(3.26) $$H_\nu H_\mu = \delta_{\nu,\mu} H_\nu,$$

we find

(3.27) $$\iint_S G(H)\, e^{-itH}\, d\tau = \sum_{\nu=1}^{n} \int_S B(\lambda_\nu)\, e^{-i\lambda_\nu t}\, H_\nu\, d\tau\ .$$

By applying Lemma 4 and Lemma 6 to (3.27) we find with $\lambda_{\nu,o} = \lambda_\nu - \sigma$,

$$(3.28) \qquad \iint_S G(H)\, e^{-itH}\, d\tau = \sum_{\nu=1}^{n} \sqrt{n} \int_{S_o} H_\nu\, d\tau' \int_{-\infty}^{\infty} d\sigma \left\{ B(\lambda_{\nu,o}+\sigma) \exp\left[-it(\lambda_{\nu,o}+\sigma)\right]\right\}$$

where S_o is the part of the space $S(0)$ of Lemma 4 that lies within the part S of S_{II} defined in the Lemmas in Section 1. It should be noted that if H_o is a point of S_o then S contains all the points $H_o + \sigma I$ for $-\infty < \quad < \infty$, and, conversely, if H is in S, then there exists a uniquely determined matrix H_o (with trace zero) in S_o and uniquely determined value of σ such that $H = H_o + \sigma I$.

By applying the ordinary Fourier theorem to (3.28) and using the identity

$$(3.29) \qquad \sum_{\nu=1}^{n} H_\nu = I ,$$

we find

$$(3.30) \qquad \int_S G(H)\, e^{itH}\, d\tau = F(t)\, 2\pi \sqrt{n} \int_{S_o} d\tau' .$$

This is the Fourier theorem (1.9) with

$$(3.31) \qquad L_n = 2\pi \sqrt{n} \int_{S_o} d\tau' = 2\pi \sqrt{n}\; V_{n,o} ,$$

where $V_{n,o}$ is the volume of S_o, computed from the volume element $d\tau'$ of Lemma 4.

The problem of computing $V_{n,o}$ seems to be a difficult one if $n > 2$. For $n = 2$, however, we can proceed in the following manner to find L_2: Let

$$(3.32) \qquad H = \begin{pmatrix} s_1 & \phi+i\Psi \\ \phi-i\Psi, & s_2 \end{pmatrix} \qquad\qquad \sigma = \frac{1}{2}\left(s_1 + s_2\right) .$$

Then

$$(3.33) \qquad \lambda_{1,2} = \sigma \pm \sqrt{\frac{1}{2}\rho_1^2 + \phi^2 + \Psi^2} , \qquad\qquad \rho_1 = \frac{s_1 - s_2}{\sqrt{2}} ,$$

and

$$(3.34) \qquad d\tau' = \frac{\sin^2(\frac{1}{2}\rho_1^2 + \phi^2 + \Psi^2)^{1/2}}{(\frac{1}{2}\rho_1^2 + \phi^2 + \Psi^2)} \, d\rho_1 \, d\phi \, d\Psi.$$

The region S_o over which we have to integrate is given by

$$(3.35) \qquad -\pi \leqq (\tfrac{1}{2}\rho_1^2 + \phi^2 + \Psi^2)^{1/2} \leqq \pi.$$

Now we can introduce polar coordinates

$$\frac{\rho_1}{\sqrt{2}} = R \cos \alpha, \qquad \phi = R \sin \alpha \cos \beta, \qquad \Psi = R \sin \alpha \sin \beta,$$

where $R^2 = \frac{\rho_1^2}{2} + \phi^2 + \Psi^2$. Therefore we have

$$\int_{S_o} d\tau' = V_{2,o} = \sqrt{2} \int_0^\pi dR \int_0^{2\pi} d\beta \int_0^\pi d\alpha \left\{ \sin^2 R \sin \alpha \right\} = \sqrt{8}\, \pi^2 .$$

This gives

$$L_2 = (2\pi)^3 .$$

4. The Plancherel Theorem

In this section we shall prove formula (1.10). We have for the left-hand side of (1.10)

$$(4.1) \qquad \sum_{\nu,\mu = 1}^{n} \int_{-\infty}^{\infty} |f_{\nu,\mu}(t)|^2 \, dt.$$

As in (3.23), we define $b_{\nu,\mu}(\lambda)$ by

$$(4.2) \qquad \int_{-\infty}^{\infty} e^{it\lambda} f_{\nu,\mu}(t)dt = b_{\nu,\mu}(\lambda) ;$$

then we find from (3.24), from $H_\nu^* = H_\nu$, and from $H_\nu H_\mu = \delta_{\nu,\mu} H_\nu$ that

$$(4.3) \qquad \text{trace } G^*G = \text{trace } G\, G^* = \sum_{\nu=1}^{n} \sum_{\ell,r,\rho=1}^{n} b_{\ell,r}(\lambda_\nu)\, h_{r,\rho}^{(\nu)}\, \bar{b}_{\ell,\rho}(\lambda_\nu) ,$$

where $h_{r,\rho}^{(\nu)}$ is the element in the r-th row and ρ-th column of H_ν.

In order to compute

$$(4.4) \qquad \iint_S \text{trace } G\, G^*\, d\tau$$

we decompose the integration [as in (3.28)] into an integration over S_o and an integration over σ from $-\infty$ to ∞. Carrying out the integration with respect to σ, and observing that $h_{r,\rho}^{(\nu)}$ is independent of σ, we find

$$(4.5) \qquad \int_{-\infty}^{\infty} b_{\ell,r}(\lambda_\nu)\, \bar{b}_{\ell,\rho}(\lambda_\nu)\, d\sigma = \int_{-\infty}^{\infty} b_{\ell,r}(\sigma)\, \bar{b}_{\ell,\rho}(\sigma) d\sigma = \gamma_{\ell,\rho,r}.$$

The $\gamma_{\ell,\rho,r}$ in (4.5) are constants which do not depend on ν, that is, they are independent of the particular eigenvalue λ_ν which appears in the left-hand side of (4.5). Because of (4.2) the ordinary Plancherel theorem gives

$$(4.6) \qquad \gamma_{\ell,r,r} = 2\pi \int_{-\infty}^{\infty} |f_{\ell,r}(t)|^2\, dt$$

for $r = \rho$. Using the fact that $H_1 + H_2 + \dots + H_n = I$, that is,

$$(4.7) \qquad \sum_{\nu=1}^{n} h_{r,\rho}^{(\nu)} = \delta_{r,\rho},$$

we find from (4.5), (4.6) and (4.7) that

$$(4.8) \qquad \int_{-\infty}^{\infty} \operatorname{trace} G\, G^* \, d\sigma = 2\pi \int_{-\infty}^{\infty} \operatorname{trace} F^*(t)\, F(t)\, dt .$$

By integrating the left-hand side of (4.8) over S_o we arrive now at (1.10).

5. Some Properties of the Transformed Functions

It is clear that the elements of a matrix $G(H)$ cannot be arbitrary functions of n^2 variables of H. We shall use Lemma 2 of Section 3 to prove that the elements of $G(H)$ can be written as derivatives of unitary invariants of H. For this purpose, we define first a unitary invariant $j(H)$ in the following manner: Let U be any unitary matrix and let $H = (s_{\nu,\mu} + ia_{\nu,\mu})$ be a hermitian matrix. Then

$$(5.1) \qquad U^{-1} H U = \hat{H} = (\hat{s}_{\nu,\mu} + \hat{i}a_{\nu,\mu})$$

is a unitary matrix again. A function

$$(5.2) \qquad j(H) = j(s_{\nu,\mu}, a_{\nu,\mu})$$

is called a unitary invariant if, for any matrix U and any variables $\hat{s}_{\nu,\mu}$, $\hat{a}_{\nu,\mu}$ derived from U by means of (5.1)

$$(5.3) \qquad j(\hat{s}_{\nu,\mu}, \hat{a}_{\nu,\mu}) = j(s_{\nu,\mu}, a_{\nu,\mu}) .$$

We state the following

Lemma 7:

A function $j(\rho_{\nu,\mu}, a_{\nu,\mu})$ is a continuously differentiable unitary invariant if and only if

$$(5.4) \qquad (\nabla_H\, j)H - H(\nabla_H\, j) = 0,$$

where ∇_H is the differential operator defined by (3.2).

The proof is based on a standard procedure. Since the space of unitary transformations is connected, a sufficient condition for a function to be a unitary invariant is that it be an invariant under infinitesimal unitary substitutions. If U is an infinitesimal unitary transformation, it can be written as

$$(5.5) \qquad U = I + i\, dK$$

where dK is an infinitesimal hermitian matrix, i.e.,

$$(5.6) \qquad \begin{cases} dK = (dp_{\nu,\mu} + i dq_{\nu,\mu}) , \\ dp_{\nu,\mu} = dp_{\mu,\nu}, \qquad\qquad dq_{\nu,\mu} = -dq_{\mu,\nu} . \end{cases}$$

In this case we have

$$(5.7) \qquad U^{-1} H U = H + i(H\, dK - dK\, H)$$

Now if $j = j(H)$ is a unitary invariant, then

$$(5.8) \qquad j(H) = j(H + i(H\, dK - dK\, H)$$

and the connection between the $s_{\nu,\mu}$, $a_{\nu,\mu}$, $\hat{s}_{\nu,\mu}$, $\hat{a}_{\nu,\mu}$ in (5.3) is given by

$$(5.9) \qquad \begin{cases} \hat{s}_{\nu,\mu} = s_{\nu,\mu} + \sum\limits_{r=1}^{n} \left(dq_{\nu,r}\, s_{r,\mu} - s_{\nu,r}\, dq_{r,\mu} + dp_{\nu,r}\, a_{r,\mu} - a_{\nu,r}\, dp_{r,\mu} \right), \\ \hat{a}_{\nu,\mu} = a_{\nu,\mu} + \sum\limits_{r=1}^{n} \left(s_{\nu,r}\, dp_{r,\mu} - dp_{\nu,r}\, s_{r,\mu} + dq_{\nu,r}\, a_{r,\mu} - a_{\nu,r}\, dq_{\nu,\mu} \right) . \end{cases}$$

If we take the first-order term of the Taylor expansion of the right-hand side of (5.8) and put the coefficients of the $dp_{\nu,\mu}$, $dq_{\nu,\mu}$ equal to zero, we find that (5.4) is a consequence of (5.6) and (5.9).

Now we can consider the elements $g_{\nu,\mu}$ of G(H). Let $b_{\nu,\mu}$ be the elements of the matrix $B(\lambda)$ defined by (3.23). Let $\hat{b}_{\nu,\mu}(\lambda)$ be the indefinite integral of

$b_{\nu,\mu}(\lambda)$, that is

(5.10) $$\frac{d\hat{b}_{\nu,\mu}}{d\lambda} = b_{\nu,\mu} .$$

Then

(5.11) $$j_{\nu,\mu}(H) = \sum_{\ell=1}^{n} b_{\nu,\mu}(\lambda_\ell)$$

is a unitary invariant of H since it is a symmetric function of its eigenvalues. From Lemma 2 of Section 3 and from (3.24) we find now

(5.12) $$g_{\nu,\mu} = \sum_{\ell=1}^{n} \left(\frac{\partial j_{\nu\ell}}{\partial s_{\ell\mu}} + i \frac{\partial j_{\nu\ell}}{\partial a_{\ell\mu}} \right) \varepsilon_{\ell\mu} ,$$

where $\varepsilon_{\nu,\mu} = \frac{1}{2}$ if $\nu \neq \mu$, and where $\varepsilon_{\nu,\nu} = 1$. Equation (5.12) is the representation of the elements of G in terms of unitary invariants of H which was mentioned in the introduction.

Reference

[1] Magnus, Wilhelm, Algebraic aspects in the theory of systems of linear differential equations, Research Report No. BR-3, New York University, Institute of Mathematical Sciences, June, 1953, or Comm. Pure Appl. Math., 7, no. 4 (1954).

(CLASSIFICATION)

Security Information

Bibliographical Control Sheet

1. Originating agency and/or monitoring agency:

O.A.: Institute of Mathematical Sciences, Division of Electromagnetic Research, New York University, New York City
M.A.: Mathematics Division, Office of Scientific Research

2. Originating agency and/or monitoring agency report number:

O.A.: Research Report No. BR-8
M.A.: OSR-TN-54-308

3. Title and classification of title:

A Fourier Theorem for Matrices (UNCLASSIFIED)

4. Personal author(s): Wilhelm Magnus

5. Date of report: October, 1954

6. Pages: i + 1 - 20

7. Illustrative material: none

8. Prepared for Contract No.(s): AF-18(600)367

9. Prepared for Project Code(s) and/or No.(s): No. R-354-10-15

10. Security classification: UNCLASSIFIED

11. Distribution limitations: none

12. Abstract:

A Fourier theorem is proved which establishes a one-to-one correspondence between $n \times n$ matrices depending on a real parameter and $n \times n$ matrices depending on the elements of a variable hermitian matrix and satisfying certain differential equations. The analogue of the Plancherel formula is shown to be true for this Fourier theorem.

NYU
BR-
8 Magnus.
c. 2
A Fourier theorem for
matrices.